

EXHIBIT B

Message

From: Aurelien Rodriguez [REDACTED]@meta.com]
Sent: 11/28/2022 5:28:15 PM
To: Melanie Kambadur [REDACTED]@meta.com]; Faisal Azhar [REDACTED]@meta.com]; Aurelien Rodriguez [REDACTED]@meta.com]; Guillaume Lample [REDACTED]@meta.com]; Susan Zhang [REDACTED]@meta.com]
Subject: Message summary [{"otherUserFbId":null,"threadFbId":5757200030967938}]

Melanie Kambadur (11/28/2022 08:49:05 PST):

Redacted - Privilege

Guillaume Lample (11/28/2022 08:53:14 PST):

Redacted - Privilege

Melanie Kambadur (11/28/2022 08:56:32 PST):

Redacted - Privilege

Guillaume Lample (11/28/2022 08:57:57 PST):

Redacted - Privilege

Guillaume Lample (11/28/2022 08:58:16 PST):

Redacted - Privilege

Guillaume Lample (11/28/2022 08:58:20 PST):

Redacted - Privilege

Melanie Kambadur (11/28/2022 08:58:43 PST):

Redacted

Guillaume Lample (11/28/2022 08:59:11 PST):

Redacted

Melanie Kambadur (11/28/2022 08:59:49 PST):

>did someone in legal confirm that? or are we just trying to not ask too many qs?

Guillaume Lample (11/28/2022 09:00:53 PST):

>I didn't ask questions :D

Guillaume Lample (11/28/2022 09:01:31 PST):

>but this is what OpenAI does with GPT3, what Google does with PALM, and what Deepmind does with chinchilla

Guillaume Lample (11/28/2022 09:01:43 PST):

>so we will do it to ^^

Melanie Kambadur (11/28/2022 09:25:48 PST):

Redacted - Privilege

Guillaume Lample (11/28/2022 09:26:31 PST):

>i think you should !

Guillaume Lample (11/28/2022 09:26:39 PST):

>also note that we are only using a fraction of libgen

Guillaume Lample (11/28/2022 09:27:12 PST):

>there are multiple splits in libgen: sci-tech (the real libgen), fiction (mostly novels, easy to parse, what we use), and sci-mag

Guillaume Lample (11/28/2022 09:27:20 PST):

>sci-mag probably bad, just PDF publications

Guillaume Lample (11/28/2022 09:27:33 PST):

>fiction is great but ~700GB only

Guillaume Lample (11/28/2022 09:27:43 PST):

>sci-tech is the real thing, will require more time to process

Melanie Kambadur (11/28/2022 09:27:47 PST):

>yeah the fiction bit is particularly interesting right since we're not in any discussions to license fiction data right now, right?

Melanie Kambadur (11/28/2022 09:27:52 PST):

>at least AFAIK

Guillaume Lample (11/28/2022 09:28:15 PST):

>not that i know of